Neural networks in gravitational-wave astronomy: The rigorous approach

Alvin Chua
JPL-Caltech

JGRG 29
Kobe University, Japan
28 November 2019

Jet Propulsion Laboratory
California Institute of Technology

Why the rigorous approach?

- Traditional approach to GW science is hard
- Forward problem: Parameters ↦ expected data
  - Template models: Time/frequency detector response
  - Noise models: PSD, transients
- Inverse problem: Data ↦ inferred parameters
  - Point estimates: MLE (overlap maximization)
  - Credible regions: Bayesian posteriors

$$\theta \mapsto h$$

$$h + n \mapsto \theta$$
$$h + n \mapsto p(\theta)$$

Why the rigorous approach?

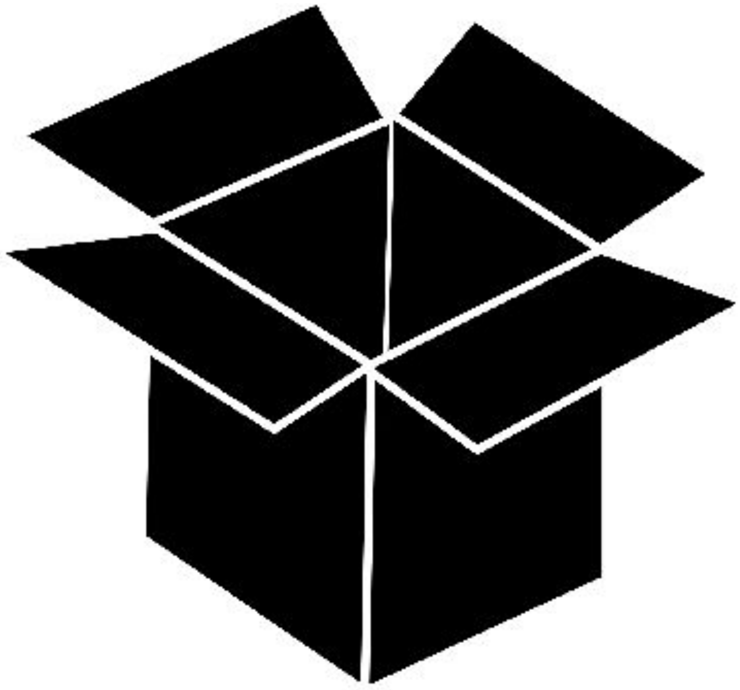
- Traditional approach to GW science is hard
- Forward problem: Parameters ↦ expected data
  - Template models: Time/frequency detector response
  - Noise models: PSD, transients
- Inverse problem: Data ↦ inferred parameters
  - Point estimates: MLE (overlap maximization)
  - Credible regions: Bayesian posteriors
- Two motivating reasons for integration
  - We want to complement & improve traditional approach, not replace it
  - We want to streamline forward/inverse solutions & their interface

$$\theta \mapsto h$$

$$h + n \mapsto \theta$$
$$h + n \mapsto p(\theta)$$

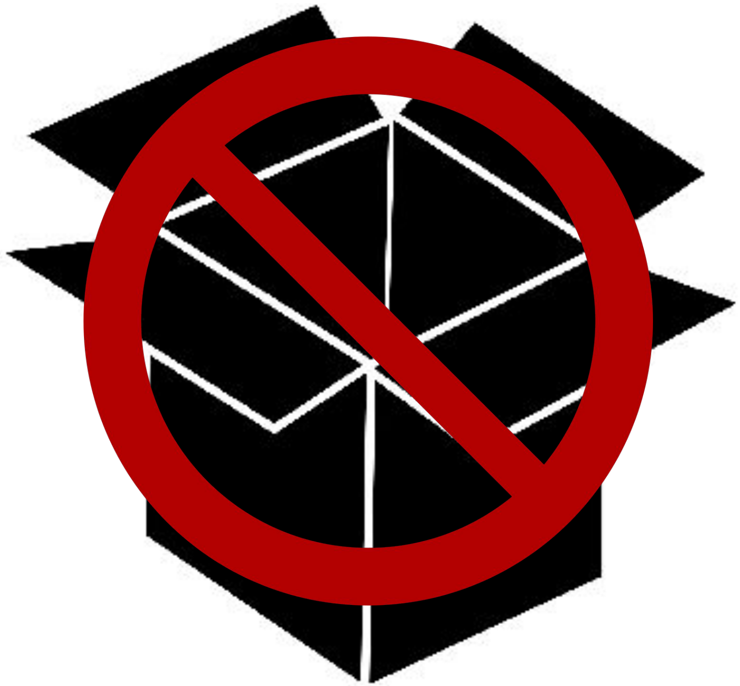
Deep learning, demystified (only for this talk)

- Highly recursive nonlinear regression
  - Input X ⟶ neural network N(X,P) ⟶ output Y
- Training: Optimize N over parameters P
- Supervised vs unsupervised
  - Training set is {(X,Y)} vs training set is {X}
- Classification vs regression
  - Y is discrete vs Y is continuous



Deep learning, demystified (only for this talk)

- Highly recursive nonlinear regression
  - Input X ⟶ neural network N(X,P) ⟶ output Y
- Training: Optimize N over parameters P
- Supervised vs unsupervised
  - Training set is {(X,Y)} vs training set is {X}
- Classification vs regression
  - Y is discrete vs Y is continuous
- Statistical model for interpolation/fitting
  - Scales well with dimensionality of X & Y
  - Y is fully analytic & fast to compute
  - dY/dX, etc. are also analytic & obtained for free



Deep learning in GW astronomy*

*Apologies if there any omissions
**No manuscript yet (I think)

- Various LIGO-type applications
  - Glitch classification (Zevin et al.; Razzano & Cuoco; George et al.)
  - Denoising (Shen et al.; Wei & Huerta; Kulkarni & Cavaglia**)
  - Detector control systems (Vajente et al.**)
- LIGO-type signal classification & regression
  - Convolutional neural networks (George & Huerta; Gebhard et al.; Gabbard et al.)
  - Follow-on analyses (Fan et al.; Rebei et al.; Nakano et al.; Field et al.**)

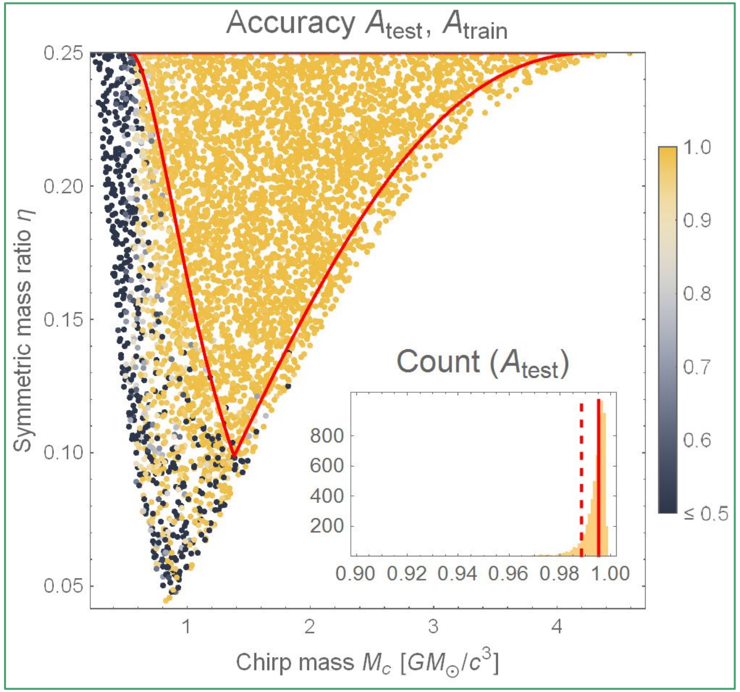
Deep learning in GW astronomy*

*Apologies if there any omissions
**No manuscript yet (I think)

- Various LIGO-type applications
  - Glitch classification (Zevin et al.; Razzano & Cuoco; George et al.)
  - Denoising (Shen et al.; Wei & Huerta; Kulkarni & Cavaglia**)
  - Detector control systems (Vajente et al.**)
- LIGO-type signal classification & regression
  - Convolutional neural networks (George & Huerta; Gebhard et al.; Gabbard et al.)
  - Follow-on analyses (Fan et al.; Rebei et al.; Nakano et al.; Field et al.**)
- Current classification work ≠ statistical signal detection
  - Test set prevalence & FPR are not representative (but see Gebhard et al. 2019)
- Current regression work ≠ Bayesian parameter estimation
  - Estimates/errors are statements about data sets, network architecture, training process
  - Have not been mapped to statements about signal model, noise model, astrophysical prior

Neural networks in forward models

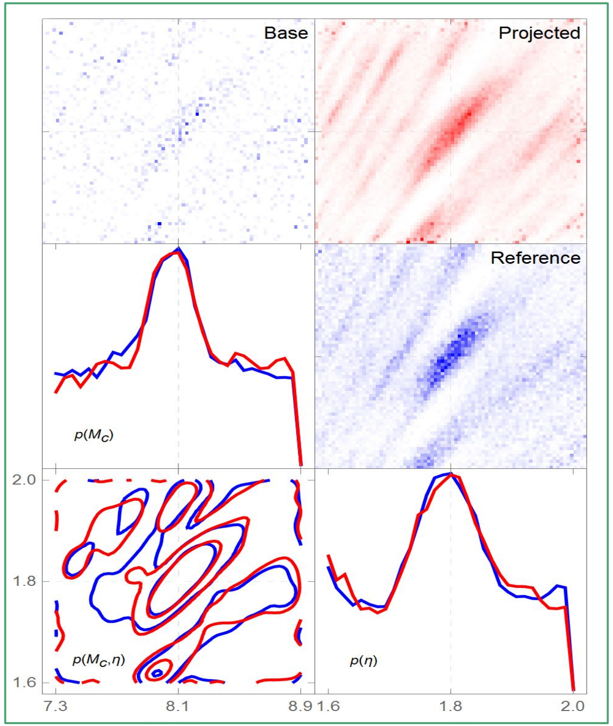
- ROMAN (Chua, Galley & Vallisneri)
  - Reduced-order modeling with artificial neurons
  - Near-lossless compression of model with ROM
  - Source parameters ↦ ROM coefficients
  - Shown on 4-parameter binary inspiral model
  - Speed/accuracy comparable to surrogates



Chua, Galley & Vallisneri (2019)

Neural networks in forward models

- ROMAN (Chua, Galley & Vallisneri)
  - Reduced-order modeling with artificial neurons
  - Near-lossless compression of model with ROM
  - Source parameters ↦ ROM coefficients
  - Shown on 4-parameter binary inspiral model
  - Speed/accuracy comparable to surrogates
- Inference can be done in ROM domain
  - Faster likelihood evaluations like ROQ
- Fast & accurate template derivatives
  - Fisher matrix estimates become trivial
  - Derivative-based sampling (MALA, HMC, etc.)
  - Derivative-based upsampling (Chua)



Chua (2019)

Neural networks as inverse models

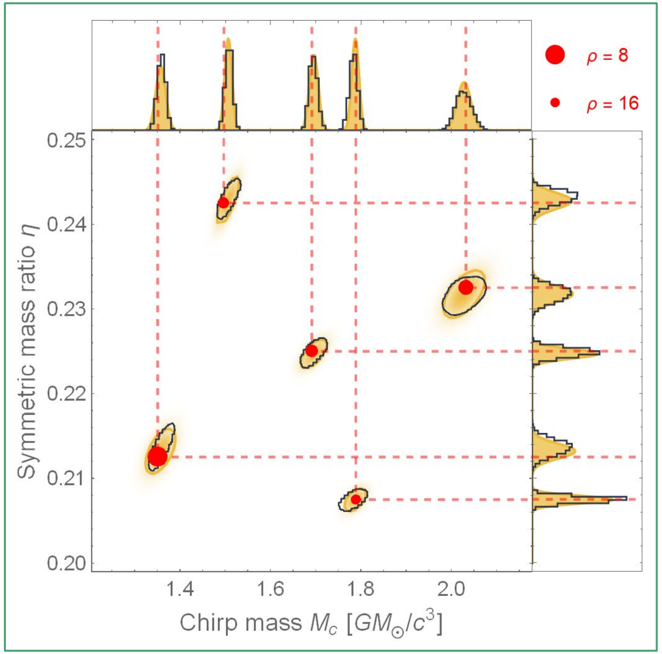
- Inverse problem without forward models
  - Difficult: Essentially parametrized by noise
  - Unlikely to remove the need for posterior sampling
  - But solvable in principle with perfect training

$$\theta(h+n)$$

$$p(\theta|h+n)$$

Neural networks as inverse models

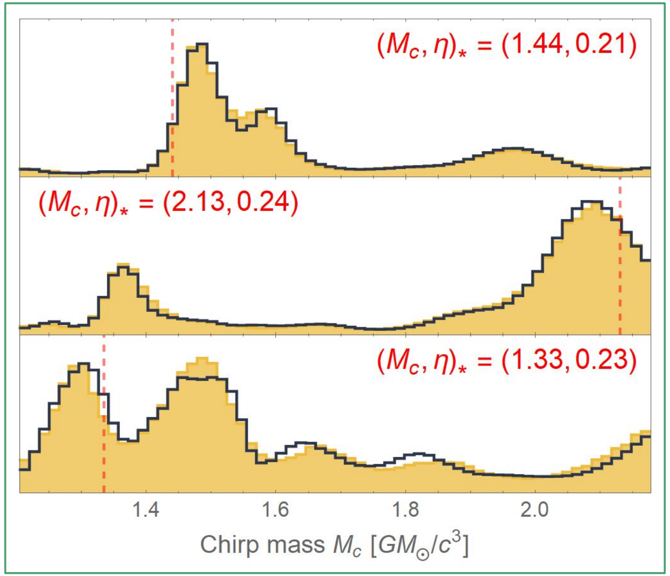
- Inverse problem without forward models
  - Difficult: Essentially parametrized by noise
  - Unlikely to remove the need for posterior sampling
  - But solvable in principle with perfect training
- PERCIVAL (Chua & Vallisneri)
  - Input: Detector data (in some representation)
  - Output: 1- or 2-parameter marginalized posteriors
  - No sampling performed during runtime
  - No posteriors computed during training
- Several potential applications
  - Fast sky localization
  - MCMC proposal kernels
  - Scoping out parameter estimation for LISA



Chua & Vallisneri (in review)

Neural networks as inverse models

- Inverse problem without forward models
  - Difficult: Essentially parametrized by noise
  - Unlikely to remove the need for posterior sampling
  - But solvable in principle with perfect training
- PERCIVAL (Chua & Vallisneri)
  - Input: Detector data (in some representation)
  - Output: 1- or 2-parameter marginalized posteriors
  - No sampling performed during runtime
  - No posteriors computed during training
- Several potential applications
  - Fast sky localization
  - MCMC proposal kernels
  - Scoping out parameter estimation for LISA



Chua & Vallisneri (in review)

Summary & references

- Deep learning is rapidly drawing interest in GW astronomy
- It is not a (completely) black box, and its limitations can be understood
- It can bring computational benefits when integrated with traditional methods

---

- A. J. K. Chua & M. Vallisneri, Learning Bayesian posteriors with neural networks for gravitational-wave inference, in review.
- A. J. K. Chua, Sampling from manifold-restricted distributions using tangent bundle projections, in press (2019).
- A. J. K. Chua, C. R. Galley & M. Vallisneri, Reduced-order modeling with artificial neurons for gravitational-wave inference, Phys. Rev. Lett. 122, 211101 (2019).